

In the Claims:

Please cancel claims 24-48 (claims 1-23 were previously canceled). Please add new claims 49-58. Set forth below, in ascending order with status identifiers, is a complete listing of all claims currently under examination. Currently amended claims are shown with additions underlined and deletions in ~~strike through~~. No new matter has been added.

1.-48. (Canceled)

49. (New) A computer implemented method of constructing a model configured to classify biological samples as being of one of at least a first state or a second state different than the first state, comprising:

providing a plurality of data strings, each data string being derived from a biological sample known to be of the first state or the second state;

using a genetic algorithm to select a first set of variables that identify data in each of the plurality of data strings;

calculating a sample vector for each member of the set of data strings using the first set of variables;

finding a location in a first vector space of each of at least two data clusters that best fit the sample vectors calculated using the first set of variables;

determining a variability for the at least two data clusters that best fit the sample vectors calculated using the first set of variables;

determining whether the variability of the at least two data clusters that best fit the sample vectors calculated using the first set of variables is within an acceptable tolerance;

if it is determined that the variability of the at least two data clusters that best fit the sample vectors calculated using the first set of variables is within the acceptable tolerance, providing the locations in the first vector space of the at least two data clusters that best fit the sample vectors calculated using the first set of variables; and

if it is determined that the variability of the at least two data clusters that best fit the sample vectors calculated using the first set of variables is not within the acceptable tolerance,

using the genetic algorithm to select a second set of variables different than the first set of variables,

calculating a sample vector for each member of the set of data strings using the second set of variables,

finding a location in a second vector space of each of at least two data clusters that best fit the sample vectors calculated using the second set of variables,

determining a variability for the at least two data clusters that best fit the sample vectors calculated using the second set of variables,

determining whether the variability for the at least two data clusters that best fit the sample vectors calculated using the second set of variables is within the acceptable tolerance, and

if it is determined that the variability of the at least two data clusters that best fit the sample vectors calculated using the second set of variables is within the acceptable tolerance, providing the locations in the second vector space of the at least two data clusters that best fit the sample vectors calculated using the second set of variables.

50. (New) The computer implemented method of claim 49, wherein the variability of the at least two data clusters that best fit the sample vectors calculated using the first set of variables is the variance of the at least two data clusters that best fit the sample vectors calculated using the first set of variables.

51. (New) The computer implemented method of claim 49, wherein if it is determined that the variability of the at least two data clusters that best fit the sample vectors calculated using the second set of variables is not within the acceptable tolerance, selecting a third set of variables different than the first set of variables and different than the second set of variables.

52. (New) The computer implemented method of claim 49, wherein each data string is derived from a biological sample via a bio-assay technique.

53. (New) The computer implemented method of claim 49, wherein the acceptable tolerance is input by a user.

54. (New) The computer implemented method of claim 49, wherein the finding a location in a first vector space of each of the at least two data clusters that best fit the sample vectors calculated using the first set of variables includes determining for each sample vector a proximity of the sample vector to a preexisting centroid in the first vector space.

55. (New) The computer implemented method of claim 54, further comprising:
determining if the distance of each sample vector from the closest preexisting centroid is within a predetermined threshold distance,
if the distance exceeds the threshold difference, defining a new centroid based on the location of the sample vector in the first vector space, and
if the distance is less than the threshold difference, assigning the sample vector to a cluster associated with the preexisting centroid.

56. (New) The computer implemented method of claim 55, wherein the assigning the sample vector includes adjusting the location of the preexisting centroid closer to the location of the sample vector.

57. (New) The computer implemented method of claim 49, wherein each of the at least two data clusters that best fit the sample vectors calculated using the first set of variables includes a centroid and a decision hyper-radius.

58. (New) A model configured to classify biological samples constructed using the method of claim 49.